УДК 336.763.3

ПОКАЗАТЕЛЬ ХЕРСТА: СПОСОБЫ РАСЧЕТЫ И ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ В ЗАДАЧАХ ПОРТФЕЛЬНОГО ИНВЕСТИРОВАНИЯ

Боровиков Илья Михайлович,
аспирант Института менеджмента, маркетинга и финансов;
ilyaflash@list.ru;
Куликова Татьяна Вячеславовна,
аспирант кафедры информационных технологий и математических методов в экономике Воронежского государственного университета; kulikovatanya08@mail.ru;
Тинякова Виктория Ивановна,
доктор экономических наук, профессор кафедры информационных технологий и математических методов в экономике Воронежского государственного университета; tviktoria@yandex.ru

Предлагается первоначальный состав портфеля ценных бумаг формировать из акций, имеющих наибольшее значение показателя Херста. Рассматриваются способы расчета этого показателя.
Ключевые слова: портфель ценных бумаг, первоначальный состав портфеля, показатель Херста.

В настоящее время инвестирование играет ключевую роль в экономических процессах. Развитие рыночной экономики и закрепление частной собственности в различных ее формах привело к тому, что наряду с денежными средствами широкое распространение в качестве средства платежа и инвестирования получили ценные бумаги.

Портфельное инвестирование можно определить как деятельность, в рамках которой инвестору приходится решать ряд задач, которые только на первый взгляд кажутся несложными – это формирование состава портфеля (т.е. выбор из огромного количества обращающихся на фондовом рынке активов, тех, которые целесообразно включить в портфель) и расчет структуры портфеля (т.е. доли каждого актива). Сущность портфельного инвестирования заключается в улучшении условий инвестирования путем формирования в определенный момент времени такой совокупности активов, которая обеспечивает новое инвестиционное качество. С учетом инвестиционных качеств ценных бумаг можно сформировать различные портфели, в каждом из которых бу¬дут соблюдены определенные пропорции между существующим риском, при-емлемым для владельца портфеля, и ожидаемым до¬ходом в определенный период времени. Проблемы формирования инвестиционного портфеля занимают одно из ведущих мест в современной экономической теории и практике, что обусловлено их актуальностью в условиях развитого рынка.

Один из основных вопросов, которые задают себе участники фондового рынка, занимающиеся инвестиционной деятельностью, – «Как определить активы, в которые следует вложить денежные средства?».

При этом необходимо учитывать, что вложения в ценные бумаги всегда сопряжены с определенным риском. Наиболее доходные ценные бумаги одновременно являются и самыми рискованными. Поэтому в экономике выработалась концепция, в соответствии с которой в целях получения оптимального результата, денежные средства должны вкладываться в различные ценные бумаги.

Комплексный подход к формированию портфеля ценных бумаг предусматривает проведение логически связанных между собой расчетов, которые условно можно разделить на 4 этапа: отбор активов; прогноз доходностей; построение портфеля и его тестирование на упреждающем периоде.

При работе на рынке ценных бумаг, инвестор должен придерживаться принципа диверсификации, т.е. стремиться к разнообразию приобретаемых финансовых активов с целью уменьшения риска потери средств.

Существуют следующие методы отбора: дивидендная доходность, использование бета-коэффициента, оценка стоимости акций компании, стратегия роста, CANSLIM, метод парных сравнений, подробно об этих методам рассказано в работе [1].

В последнее время появляются работы посвященные построению портфеля ценных бумаг на основе прогнозных значений их доходностей, а проблеме первичного отбора должного внимания не уделяется.

Важно найти такой критерий отбора, который позволит отбирать активы, поведение которых в дальнейшем может быть предсказано. Таким критерий будем считать показатель Херста. Для его оценки существует много методов, рассмотрим их более подробно.

Методы оценки параметра самоподобия (показателя Херста) временного ряда

Динамика сложных систем часто принадлежит дробному (фрактальному) броуновскому шуму, например: объём стока рек, урожайность сельскохозяйственных культур, цены активов на финансовом рынке и т.д., который описывается показателем Херста. Отсюда, востребованность инструментария для анализа таких шумов.

На финансовом рынке оценка показателя Херста необходима, например, для решения следующих задач:

- выбора типа индикаторов, сигналами которых следует руководствоваться в данный момент торговой системе: либо трендовые (если H>0,5), либо контртрендовые индикаторы (если H<0,5);

- повышения эффективности торговой стратегии за счёт введения обратной связи с результативностью предыдущих сделок;

- определение горизонта прогнозируемости;

- управления риском и др.

1. Определение понятия самоподобного временного ряда (процесса).

Самоподобные процессы были впервые рассмотрены Колмогоровым в работе 1941 [14]. Они были освещены Мандельбротом и его соавторами в конце 1960 и 1970 годах [15 – 21] и в др. работах.

Традиционное определение самоподобного процесса $X = X_t : t \geq 0$ с нулевым средним, стационарными приращениями и параметром самоподобия H (показатель Херста) даётся, как равенство многомерных распределений X_{at} и $a^H X_t$, при всех $a > 0$ [13].

Однако данное определение не удобно с точки зрения анализа временных рядов и оценки параметра Херста [23], поэтому перейдём к альтернативному определению, которое использует агрегированный процесс.

Рассмотрим ковариантно стационарный (в широком смысле) процесс $X = X_t : t = 0,1,2,...$. Обозначим: среднее – μ , дисперсию – σ^2, автокорреляционную функцию – $\rho(k),\ k = 0,1,2,...$. Перейдём от ряда X_t к агрегированному ряду $X_k^{(m)}$, который также будет является ковариантно стационарным процессом, путём агрегации X_t на n непересекающихся интервалов по $m = 1,2,...$ точек, номер интервала $k = 1,...,n$, $n \approx N/m$, N – число точек исходного ряда, k и n – целые числа:

$$X_k^{(m)} = \frac{1}{m}\sum_{t=km-m+1}^{km} X_t\ , \tag{1}$$

В зависимости от соотношения свойств агрегированного $X_k^{(m)}$ и исходного X_t процесса выделяют три типа самоподобных процессов, с параметром самоподобия H [23]:

a) процесс X_t называется строго самоподобным, если для всех натуральных $m = 1,2,...,$ X_t и $m^{-H}\sum_{t=km-m+1}^{km} X_t$ имеют одинаковое многомерное распределение;

b) процесс X_t называется самоподобным в точности второго порядка, если дисперсия и автокорреляционная функция X_t равна дисперсии и автокорреляционной функции $m^{-H}\sum_{t=km-m+1}^{km} X_t$. Таким образом, дисперсия агрегированного самоподобного процесса выражается функцией $\mathrm{var}(X_k^m) = \sigma^2 m^{2H-2}$, а автокорреляционные функции этих процессов равны и выражаются [23]:

$$\rho_k^{(m)} = \rho_k = 0,5\delta^2(|k|^{2H}) = 0,5(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}).$$

Примером самоподобного процесса, с параметром является фрактальный гауссов шум [22], введённый в работе [19], или, что одно и то же, приращение фрактального броуновского движения;

c) процесс X_t называется асимптотически самоподобным в точности второго порядка, если: $\rho_k^{(m)} \to 0,5\delta^2(|k|^{2H}),\ m \to \infty$, $\rho(k),\ k = 0,1,2,\dots$.

Примером асимптотически самоподобного процесса точности второго порядка является модель $ARFIMA(p,d,q)$ (другое название – $FARIMA$) с параметром $0 < d < 0,5$, причём $d = H - 0,5$ [23 – 24].

2. Определение понятия долговременной зависимости временного ряда.

Определение понятия процесса с долговременной зависимостью во временной области даётся на основе автокорреляционной функции. Временной ряд, обладающий свойством слабой формы стационарности, называется рядом с долговременной зависимостью, если его автокорреляционная функция $\rho(k)$ является степенной функцией (гиперболической) от k и имеет вид [31]:

$$\rho(k) \underset{k\to\infty}{\sim} \frac{L(k)}{k^{1-2d}}, \tag{2}$$

где $0 < d < 1/2$, $d = 2 - 2H$, L – функция, медленно изменяющаяся у бесконечности: $\frac{L(at)}{L(t)} \to 1$ при $t \to \infty$.

Следствием (2), для процесса с долговременной зависимостью, является расходимость ряда автокорреляционной функции $\rho(k)$, так $\sum_{k=-\infty}^{\infty} \rho(k) = \infty$ [5].

В противоположность этому, ряд с кратковременной зависимостью имеет автокорреляционную функцию:

$$|\rho(k)| \le Kc^k,\ \ K > 0,\ c \in [0,1]. \tag{3}$$

Определение долговременной зависимости в спектральной области даётся на основе спектральной плотности. Временной ряд, обладающий свойством слабой стационарности, имеет долговременную зависимость, если выполняется:

$$f(\lambda) \sim C_f |\lambda|^{-\beta},\ \ \lambda \to 0. \tag{4}$$

где $C_f > 0$, $\beta \in (0,1)$, $H = (1+\beta)/2$, $f(\lambda) = \frac{\sigma^2}{2\pi} \sum_{k=-\infty}^{\infty} p(k)e^{ik\lambda}$ – спектральная плотность, $i^2 = -1$.

Связь между процессами с долговременной зависимостью и самоподобными процессами устанавливается через спектральную плотность: $C_H |\lambda|^{1-2H}$ при $0,5 < H < 1$ [20].

Таким образом, долговременная зависимость приращений, в условиях стационарных приращений ряда, синонимична свойству самоподобия ряда, которое описывается показателем Херста – H, когда $H \in (1/2,1)$ [5].

3. Свойства и взаимосвязь процессов с долговременной зависимостью и самоподобных процессов.

Манделброт и Несс [21] показали, что самоподобные процессы и процессы с долговременной зависимостью связаны через спектральную плотность: $f(\lambda) \sim a|\lambda|^{1-2H}$, при $\lambda \to 0$, на участке $0 < H < 0,5$ где $f(\lambda) = \sum_k r(k)e^{ik\lambda}$ – спектральная плотность, $i^2 = -1$.

Типичным примером самоподобного процесса, приращения которого имеют долговременную зависимость, является фрактальное броуновское движение. Причём свойство самоподобия существует на участке $0 < H < 1$, а долговременная зависимость приращений на участке $0,5 < H < 1$ [32].

Уже из неравенства области самоподобия и долговременной зависимости фрактального броуновского движения ясно, что наличие свойства самоподобия процесса не является достаточным условием наличия долговременной памяти приращений. В случае альфа стабильных процессов самоподобие возникает на фоне полной некоррелированности приращений, за счёт особой формы распределения с тяжёлыми хвостами. Верно и обратное, наличие долговременной зависимости в приращениях не является достаточным условием свойства самоподобия процесса, примеры таких процессов с гауссовым распределением приведены в работе [48].

Отражения взаимосвязи различных процессов представлено на рисунке.

Рис. Взаимосвязь самоподобных процессов с процессами Леви и Гауссовых процессов [32]

Самоподобие может быть следствием двух причин: долговременной коррелированности (долговременной зависимости) приращений и/или распределения с тяжёлыми хвостами. Мандельброт назвал первое из них эффектом Джозефа, а второй – эффектом Ноя [49].

Ввиду того, что показатель Херста характеризует самоподобные процессы, а также процессы с долговременной памятью, но в общем случае эти процессы не синонимичны, то необходимо разделять свойства этих двух процессов (см. табл. 1).

Таблица 1

Свойства процессов

Свойства самоподобных процессов	Свойства процессов с долговременной памятью
Эффект Херста – ожидание нормированного размаха описывается степенным законом: $E[R(m)/S(m)] \sim a_1 m^H$ при $m \to \infty$. Угасание вариации – дисперсия выборочного среднего снижается как: $Var(X_k^{(m)}) \sim a_2 m^{2H-2}$ при $m \to \infty$. Спектральная плотность подчиняется степенному закону: $f(\lambda) \sim a_3 \lvert\lambda\rvert^{1-2H}$, при $\lambda \to 0$, $0 < H < 1$, где $f(\lambda) = \sum_k r(k) e^{ik\lambda}$ – спектральная плотность.	Долговременная зависимость – снижение автокорреляционной функции подчиняется не геометрическому, а гипербрлическому закону что ведёт к бесконечной сумме: $\sum_k r(k) = \infty$. Поэтому, несмотря на то, что отдельные значения $r(k)$ при больших k, могут быть малы, их суммарный эффект оказывается существенным. Спектральная плотность подчиняется степенному закону: $f(\lambda) \sim a_3 \lvert\lambda\rvert^{1-2H}$, при $\lambda \to 0$, $0{,}5 < H < 1$ где $f(\lambda) = \sum_k r(k) e^{ik\lambda}$ – спектральная плотность.

Величины a_1, a_2, a_3 – конечные, положительные константы.

Ввиду наличия взаимосвязи между процессами с долговременной памятью и самоподобными процессами, можно сформулировать следующие задачи по анализу временных рядов:

- оценка показателя самоподобия H;

- определение класса, к которому относится временной ряд: а) ряд с долговременной зависимостью без самоподобия; б) самоподобный ряд, но без долговременных зависимостей в приращениях; в) ряд, содержащий как долговременную зависимость приращений, так и являющийся самоподобным;

- дифференциация вклада в показатель H эффекта Джозефа (долговременной зависимости) и эффекта Ноя;

- исследование свойств ряда знаков и амплитуд процесса в зависимости от параметра самоподобия.

В данной статье рассматриваются методы решения первой из поставленных задач.

Необходимо также отметить, что на сегодняшний день, получила развитие концепция мультифрактального анализа, которая предполагает эволюцию параметра самоподобия при изменении масштаба. Однако такого рода усложнение не меняет базовых задач и подходов к анализу.

4. Методы оценки параметра самоподобия (показателя Херста) временного ряда.

Оценка параметра самоподобия (далее будем называть показателем

Херста) возможна на основе того или иного свойства самоподобного процесса [12]:

• Анализ во временной области, (включает $\mathrm{R/S}$-анализ).

• Анализ дисперсии агрегированного процесса, (ведёт к оценке на основе зависимости дисперсии от параметра агрегации m).

• Анализ в частотной области: аппроксимация Уайтла метода максимального правдоподобия и др. [29].

Рассмотрим семь основных методов оценки показателя Херста (H), каждый из которых основан на том или ином свойстве самоподобного процесса.

R/S-анализ.

Автором этого метода является британский гидролог Херст, который опубликовал его в своей работе [33] в 1951 году.

Значительный вклад в развитие и популяризацию метода внёс Мандельброт и его соавторы [21, 19, 34], а так же работы других авторов [31, 35].

Оценка показателя Херста посредством $\mathrm{R/S}$-статистики основывается на соотношении:

$$M[R(\tau)/S(\tau)] \sim (c\tau)^H . \tag{5}$$

Выражение (5), называют законом Херста. Преобразовав его, получим:

$$\ln\{M[R(\tau)/S(\tau)]\} \sim \ln(c) + H\ln(\tau) . \tag{6}$$

Величина $\ln(c)$ является константой. Таким образом, показатель Херста равен тангенсу угла наклона точек на графике $A(\tau) = \ln\{R(\tau)/S(\tau)\}$ от $B(\tau) = \ln(\tau)$, который обычно находят линейной регрессией по МНК:

$$H = (B^T B)^{-1} B^T A , \tag{7}$$

где вектор независимой переменной B должен быть центрирован.

Величина нормированного размаха может быть определена формулой:

$$R(\tau)/S(\tau) = \frac{1}{N}\sum_{i=1}^{N} \frac{\max\limits_{1\le t\le\tau}\left(\sum\limits_{j=1}^{t} x_{j,i} - \overline{x}_i\right) - \min\limits_{1\le t\le\tau}\left(\sum\limits_{j=1}^{t} x_{j,i} - \overline{x}_i\right)}{\sqrt{\frac{1}{\tau}\sum\limits_{j=1}^{\tau}(x_{j,i} - \overline{x}_i)^2}} ,\quad t = 1,\dots,\tau , \tag{8}$$

где τ – длина интервала разбиения; N – количество интервалов разбиения по τ значений, $N\tau = n$; $\overline{x}_i$ – среднее по i-му интервалу длины τ; $x_{j,i}$ – j-е значение x в i-м интервале разбиения, включающем τ значений Y.

Для снижения смещённости оценки H отбрасывается некоторое число точек при малых τ. В то же время для снижения дисперсии оценки H отбрасывают некоторое число точек при больших τ.

К достоинствам метода можно отнести устойчивость оценки относительно формы распределения, в том числе несимметричных распределений и распределений с длинными хвостами [12, 22].

Недостатками метода, на основе $\mathrm{R/S}$-анализа, считается отсутствие анализа надёжности и устойчивости к ошибкам [12]. Кроме того, для гауссовых моделей метод $\mathrm{R/S}$ теряет свои преимущества скажем в сравнении с оценками на основе метода максимального правдоподобия в аппроксимации Уайтла [22] а так же в случае не стационарных рядов.

Для повышения надёжности метода отбрасывают некоторое число значений при крайних низких и высоких τ [36].

Ло модифицировал метод R/S-анализа, изменив величину нормировки со стандартного отклонения на взвешенную автоковариационную функцию. Его модификация не предполагает разбиение на множество интервалов [37]. В этой модификации вместо стандартного отклонения используется показатель:

$$\hat{S}_N^2(q) = \frac{1}{N}\sum_{j=1}^{N}(X_j - \bar{X}_N)^2 + \frac{2}{N}\sum_{j=1}^{q}\omega_j(q)\left(\sum_{i=j+1}^{N}(X_i - \bar{X}_N)(X_{i-j} - \bar{X}_N)\right) =$$

$$= S_N^2 + 2\sum_{j=1}^{q}\omega_j(q)\rho_j\,, \tag{9}$$

где $\omega_j(q) = 1 - \dfrac{j}{q-1}$, $q < n$, ρ_j – оценка автоковариационной функции.

Таким образом, модифицированный показатель:

$$Q_n = \frac{R_N}{\hat{S}_N(q)}\,. \tag{10}$$

Ло показал, что распределение статистики:

$$V_N(q) = N^{-1/2} R_N / \hat{S}_N(q) \tag{11}$$

для ряда с кратковременной зависимостью имеет конечную положительную величину кумулятивного распределения [38]:

$$P(V) = 1 + 2\sum_{k=1}^{\infty}(1 - 4k^2V^2)\exp(k^2V^2)\,. \tag{12}$$

Может быть построен доверительный интервал для проверки нулевой гипотезы отсутствия долговременной зависимости – с вероятностью 95% интервал [0,809; 1,862]:

$$\lim_{N\to\infty} P\{V_N(q) \in [0,809; 1,862]\} = 0,95\,. \tag{13}$$

Когда временной ряд имеет долговременную память, величина R_N является так называемым броуновским мостом, что позволяет построить доверительный интервал. Можем классифицировать временной ряд как ряд с долговременной или кратковременной зависимостью, с помощью формулы, при больших q [38]:

$$V_N(q) = N^{-1/2} Q_n = \begin{cases} 0, & H \in (0; 0,5); \\ \infty, & H \in (0,5; 1). \end{cases} \tag{14}$$

Исходный метод Ло лишь позволяет проверить гипотезу наличия долговременной зависимости, но впоследствии в работе [40], этот метод был развит для оценки H. Было предположено использовать V_q вместе с q, строя зависимость на плоскости, как это делается в исходной R/S-статистике.

Метод оценки H на основе зависимости дисперсии от уровня агрегации ряда самоподобного процесса (метод вариограммы).

Данный метод оценки H обсуждается среди других методов в работах: [7, 37].

Этот метод основан на свойстве самоподобных процессов $Var(X_k^{(m)}) \sim am^{2H-2}$ при $m \to \infty$, где a – конечная положительная константа.

Описание алгоритма оценки H в соответствии с данным методом начинается с формирования агрегаций исходного ряда по формуле:

$$\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})} = \frac{1}{m}\sum\nolimits_{t=km-m+1}^{km} x(t), \quad \mathrm{k} = 1,2,\ldots,\mathrm{N/m}, \tag{15}$$

где $\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})}$ – агрегированный ряд из k точек, каждая из которых получена суммированием m соседних точек исходного ряда; N – число значений исходного ряда.

Для различных натуральных m из интервала от 1 до $\mathrm{N}/2$ можем рассчитать дисперсию агрегированного ряда $\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})}$:

$$D_m = \mathrm{var}(\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})}) = \frac{1}{N/m-1}\sum_{k=1}^{N/m}(\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})} - \overline{\mathrm{z}}^{(\mathrm{m})})^2, \tag{16}$$

которая будет связана с параметром масштаба агрегации по формуле:

$$D_m \sim am^{2H-2}. \tag{17}$$

Тангенс угла наклона b, линейной регрессии точек в двухмерном пространстве $\mathrm{A} = \ln(D_m)$ от $B = \ln(m)$ связан с H функцией: $b = 2H - 2$, откуда $\mathrm{H} = \mathrm{b}/2 + 1$, поэтому, используя матричную запись формулы наклона линейной регрессии по МНК, получаем оценку H:

$$\hat{H} = (B^T B)^{-1} B^T A / 2 + 1, \tag{18}$$

где вектор независимой переменной B должен быть центрирован.

Для устойчивости решения следует отбросить некоторое число точек при малых и больших m.

Этот метод затрудняет определение качества и устойчивости оценки H.

Метод оценки H на основе зависимости величины ожидаемых модулей отклонений от уровня агрегации ряда самоподобного процесса.

Алгоритм схож с предыдущим методом дисперсии агрегированного ряда и является его обобщением. В нём ряды с различными уровнями агрегации $\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})}$ так же получены с помощью формулы (15), но вместо дисперсии от уровня агрегации вычисляется ожидаемый модуль отклонения, возведённый в степень q:

$$M_m = \frac{1}{N/m-1}\sum_{k=1}^{N/m}\left|\mathrm{z}_{\mathrm{k}}^{(\mathrm{m})} - \overline{\mathrm{z}}^{(\mathrm{m})}\right|^q. \tag{19}$$

Величина M_m асимптотически равна $m^{q(H-1)}$. Поэтому тангенс угла наклона b, линейной регрессии точек в двухмерном пространстве $\mathrm{A} = \ln(M_m)$ от $B = \ln(m)$ связан с H функцией: $b = q(H-1)$, откуда

$\mathrm{H} = \mathrm{b}/\mathrm{q} + 1$, поэтому, используя матричную запись формулы наклона линейной регрессии по МНК, получаем оценку H:

$$\hat{H} = (B^T B)^{-1} B^T A / q + 1, \tag{20}$$

где вектор независимой переменной B должен быть центрирован.

Некоторое число крайних точек с обоих концов не используются для повышения устойчивости оценки.

Существует также модификация данного метода, для частного случая $q = 1$, предложенная Хигучи, состоящая в использовании при агрегации вместо непересекающихся окон, скользящего окна [76]. Алгоритм данной модификации состоит в расчёте статистики:

$$L_m = \frac{N-1}{m^3} \sum_{i=1}^{m} \frac{m}{N-1} \sum_{k=1}^{(N-i)/m} \left| \sum_{j=mk-m+i+1}^{i+km} x(j) \right|. \tag{21}$$

Величина L_m асимптотически равна cm^{H-2}. Поэтому тангенс угла наклона b, линейной регрессии точек в двухмерном пространстве $\mathrm{A} = \ln(L_m)$ от $B = \ln(m)$ связан с H функцией: $b = H - 2$, откуда $\mathrm{H} = b + 2$, поэтому, используя матричную запись формулы наклона линейной регрессии по МНК, получаем оценку H:

$$\hat{H} = (B^T B)^{-1} B^T A + 2, \tag{22}$$

где вектор независимой переменной В должен быть центрирован.

По сравнению с методом, использующим агрегацию на непересекающихся интервалах, данный метод более точный, но требует значительно большее число вычислительных операций.

Метод оценки H на основе дисперсии детрендированных остатков.

Считается, что данный метод предложен Пенгом (Peng) и соавторами, описывается, например, в обзоре [37].

Алгоритм данного метода состоит в разделении исходного ряда x, состоящего из N значений, на K интервалов по m значений обозначим $z_k^{(m)}(i)$ и последующем подсчёте векторов $y_k^{(m)}(i)$ для каждого интервала:

$$y_k^{(m)}(i) = \sum_{i=1}^{N/m} z_k^m(i),\ k = 1,\ldots,N/m,\ \ m = 2,\ldots,N/2, \tag{23}$$

где k – номер интервала; i – номер значения в интервале; m – число точек в интервале.

Далее из каждого интервала вычитается локальный (для данного интервала) линейный тренд $\hat{y}_k^{(m)}$, определяемый по МНК, в результате $Y_k^{(m)}(i) = y_k^{(m)}(i) - \hat{y}_k^{(m)}(i)$. На основе $Y_k^{(m)}$ находим среднюю сумму квадратов для всех интервалов отдельно для различных m:

$$D_m = \frac{1}{N/m} \sum_{k=1}^{N/m} \left(\frac{1}{m} \sum_{i=1}^{m} [Y_k^{(m)}(i)]^2 \right) = \frac{1}{N} \sum_{k=1}^{N/m} \sum_{i=1}^{m} [Y_k^{(m)}(i)]^2. \tag{24}$$

Аналогично базовому методу, здесь также существует закон $D_m \sim m^{2H}$.

Поэтому тангенс угла наклона b, линейной регрессии точек в двухмерном пространстве $\mathrm{A} = \ln(D_m)$ от $B = \ln(m)$ связан с H функцией: $b = 2H$, откуда $\mathrm{H} = b/2$, поэтому, используя матричную запись формулы наклона линейной регрессии по МНК, получаем оценку H:

$$\hat{H} = (B^T B)^{-1} B^T A / 2, \tag{25}$$

где вектор независимой переменной B должен быть центрирован.

Метод оценки H на основе периодограммы.

Метод оценки H на основе наклона спектральной плотности, так же называемый метод Даниэля (Daniell), имеет в качестве наиболее цитируемых первоисточников [47, 29]. Также данный метод рассматривается в работах [6, 22, 36, 37] и др.

Алгоритм метода периодограмм состоит в расчёте квадрата спектральной плотности (периодограммы):

$$I_N(\omega) = \frac{1}{2\pi N} \left| \sum_{t=1}^{N} X_j e^{it\omega} \right|^2, \tag{26}$$

где ω – частота $\omega_j = 2\pi j / N$, $j = 1 \ldots T-1$; N – число значений временного ряда X; $I_N(\omega)$ – периодограмма; $i^2 = -1$.

Величина $I_N(\omega)$, для ряда с долговременной зависимостью, пропорциональна $|\omega|^{1-2H}$. Поэтому тангенс угла наклона b, линейной регрессии точек в двухмерном пространстве $\mathrm{A} = \ln[I_N(\omega)]$ от $B = \ln(\omega)$ связан с H функцией: $b = 1 - 2H$, откуда $\mathrm{H} = 0,5 - b/2$, поэтому, используя матричную запись формулы наклона линейной регрессии по МНК, получаем оценку H:

$$\hat{H} = -(B^T B)^{-1} B^T A / 2 + 0,5, \tag{27}$$

где вектор независимой переменной B должен быть центрирован.

На практике при оценке H по данному методу, для повышения точности оценки, используют только первые 10% частот (10% самых низких частот).

Метод оценки H на основе аппроксимации Уиттла периодограммы.

Идея данного метода была предложена в [26], а развита в работах [27, 28, 31]. Данный метод оценки является модификацией метода периодограмм для гауссовых процессов [36].

Алгоритм этого метода предполагает расчёт периодограммы $I(\omega)$ по формуле (26). Далее составляется функция:

$$L_W(\vartheta) = \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} \ln[f(\omega;\vartheta)] d\omega + \int_{-\pi}^{\pi} \frac{I(\omega)}{f(\omega;\vartheta)} d\omega \right), \tag{28}$$

где $\vartheta = (\vartheta_1, \eta)$ – вектор параметров; $I_N(\omega)$ – периодограмма; $f(\omega, \eta)$ – параметрическая спектральная плотность у частоты; ω – частота.

Введём переменные: $\vartheta^* = (1, \eta)$, $\eta = (\vartheta_2, \vartheta_3, \ldots, \vartheta_M)$, $\vartheta_1 = 2\pi\sigma_\varepsilon^2$, σ_ε^2 – ошибка оптимального прогноза на 1 шаг, т.е. дисперсия случайной компоненты (инновации) авторегрессии бесконечного порядка, $\sigma_\varepsilon^2 = 2\pi\hat{\vartheta}^*$.

Дискретная аппроксимация (28) задаётся [22]:

$$\tilde{L}_W(\vartheta) = \frac{2}{m}\left(\sum_{j=1}^{m^*}\ln[f(\omega;\vartheta)] + \sum_{j=1}^{m^*}\frac{I(\omega)}{f(\omega;\vartheta)}\right), \qquad (29)$$

где $\omega = 2\pi j/m,\ j = 1,\dots,m^*,\ m^* = (m-1)/2$ принимает только целые значения.

Использование алгоритма быстрого преобразования Фурье позволяет вычислять $I(\omega)$ очень быстро. Более того, если выполняются условия

$$f(\omega;\vartheta) = \vartheta_1 f(\omega;\vartheta^*) \text{ и } \int_{-\pi}^{\pi}\ln[f(\omega;\vartheta^*)]d\omega = 0, \qquad (30)$$

то можем заменить первый интеграл на $\ln(\vartheta_1) = \int_{-\pi}^{\pi}\ln[f(\omega;\vartheta)]d\omega$. В таком случае минимизация $L_W(\vartheta)$ за счёт подбора вектора ϑ, будет эквивалентна минимизации суммы коэффициентов $I(\omega)f(\omega;\vartheta)^{-1}$. Значит, в дискретном приближении требуется минимизировать функцию, подобрав вектор η [22]:

$$\tilde{Q}(\eta) = \sum_{j=1}^{m}\frac{I(\omega)}{f(\omega;\vartheta^*)} \to \min. \qquad (31)$$

Оценка параметра σ_ε^2 выполняется по формуле:

$$\hat{\sigma}_\varepsilon^2 = 2\pi\hat{\vartheta}_1 = \frac{4\pi}{m^*}\tilde{Q}(\hat{\eta}). \qquad (32)$$

Эта дискретная аппроксимация для фрактального броуновского шума предложена в [41].

Доверительный интервал $\hat{H}$ задаётся:

$$\hat{H} \pm z_\alpha\sqrt{V/N}, \qquad (33)$$

где V – матрица $V = 2D^{-1}$ и D матрица:

$$D_{i,j} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta_i}\ln f(\lambda)\frac{\partial}{\partial\theta_j}\ln f(\lambda)d\lambda. \qquad (34)$$

В случае подгонки фрактальным броуновским шумом переменные ϑ и $f(\omega;\vartheta)$ определяются как [22]:

$$\vartheta = [\sigma_\varepsilon^2/(2\pi), H], \qquad (35)$$

$$f(\omega;\vartheta) = \sigma_\varepsilon^2/(2\pi)[c(H)f_1(\omega;H)], \qquad (36)$$

где $f_1(\omega;H) = \sigma^2/\pi\Gamma(2H+1)\sin(\pi H)(1-\cos(\omega))\sum_{j=-\infty}^{\infty}\left|2\pi j+\omega\right|^{-2H-1}$; $\sigma^2 = \operatorname{var}(X_t)$; $c(H) = \exp\left(-\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln(f_1(\omega))d\omega\right)$.

В случае описания моделью *FARIMA* (*0,d,0*):

$$f(\omega;\vartheta) = \sigma_\varepsilon^2/(2\pi)\left|1-e^{i\omega}\right|^{1-2H}, \qquad (37)$$

$$f_1(\omega;H) = \left|1-e^{i\omega}\right|^{1-2H}. \qquad (38)$$

Одним из важных преимуществ метода аппроксимации Уайтла является возможность построения доверительного интервала.

К недостаткам метода относится ограниченность использования классом гауссовых процессов, так как именно в этом случае спектральная плотность полностью характеризует процесс.

Другие методы, основанные на периодограмме, рассматриваются в работах [29, 30].

Метод оценки H на основе вейвлет-анализа.

Данный метод развивается во многих работах, так как в настоящее время техника вейвлет-анализа получила значительную популярность. Наиболее цитируемыми авторами являются Абри (Abry) [44],[40], Селан (Sellan), Мейер (Meyer) и др. Новый тип вейвлета, который наиболее успешно используется для оценки H, предложен Добеши (Daubechies) [45].

Идея оценки H на основе вейвлет анализа можно назвать модификацией метода периодограмм, так как в обоих случаях используется свойство спектральной плотности самоподобного процесса $\Gamma = 1/f^{\gamma}$ (для фрактального гауссового шума $\gamma = 2H - 1$) [43], которая в случае вейвлет-оценки представлена несколько иначе – в виде функции от вейвлет-коэффициентов.

Отличием данного метода от метода периодограмм является возможность использовать иные базисы разложения, помимо тригонометрических функций, а так же возможность разрешения во времени.

Данный метод оценки H начинается с оценки вейвлет-коэффициентов:

$$\psi_{j,k}(t) = |a|^{-m/2}\,\psi(a^{-j}t - k),\ m,k \in I\,, \tag{39}$$

$$d_x(j,k) = \int_{-\infty}^{\infty} x_t \psi_{j,k}(t)dt\,. \tag{40}$$

На практике, как правило, параметр a=2, так как для этого случая разработан алгоритм быстрого вейвлет-преобразования. В этом случае формула вейвлет-коэффициентов:

$$d_x(j,k) = 2^{-m/2}\int_{-\infty}^{\infty} x_t \psi(2^{-j}t - k)dt\,, \tag{41}$$

где j – параметр масштаба (уровня); k – номер коэффициента на уровне j; ψ – базовая вейвлет-функция; x_t – временной ряд.

Для оценки H на основе вейвлет-коэффициентов $d_x(j,k)$, наиболее используемым является метод предложенный Абри, например, [40], который заключается в расчёте функции:

$$\mu_j = \frac{1}{N(j)}\sum_{k=1}^{N(j)} |d_x(j,k)|^2\,. \tag{42}$$

Функция (42) соотносится с показателем Херста H выражением:

$$\mu_j \sim 2^{\gamma \times j + c}\,, \tag{43}$$

где γ – функция показателя Херста, причём для фрактального гауссового шума $\gamma = 2H - 1$, $c = 0$.

Прологарифмировав по основанию 2, получаем метод практической оценки H:

$$\log_2(\mu_j) \sim (2H-1)j + c\,. \tag{44}$$

Откуда показатель H может быть выражен, как функция от тангенса угла наклона (b) вида $b = 2H - 1$, с помощью линейной регрессии МНК, аппроксимирующей точки в двухмерном пространстве от $A = \log_2(\mu_j)$. Таким образом, показатель Херста: $H = b/2 + 0,5$ или с помощью матричной формулы:

$$\hat{H} = -(B^T B)^{-1} B^T A / 2 + 0,5\,, \tag{45}$$

где вектор независимой переменной B должен быть центрирован.

В работе [44] предлагается следующая формула в явном виде для оценки H с помощью вейвлет-разложения, в которой используется только некоторый интервал уровней вейвлет-коэффициентов:

$$\hat{H}(j_1, j_2) \equiv \frac{1}{2}\left[\frac{\sum_{j=j1}^{j2} jS_j\eta_j - \sum_{j=j1}^{j2} jS_j \sum_{j=j1}^{j2} S_j\eta_j}{\sum_{j=j1}^{j2} S_j \sum_{j=j1}^{j2} S_j j^2 - \left(\sum_{j=j1}^{j2} S_j j\right)^2} + 1\right], \quad (46)$$

где j_1, j_2 – масштабы (уровни) между которыми проводится оценка (когда переменная стоит в индексе другой переменной то, для удобства j_1, j_2 будем записывать как $j1, j2$), $S_j = (n\ln^2 2)/2^{j+1}$ – обратная теоретическая асимптотическая вариация η_j; $d_x(j,k)$ – вейвлет-коэффициенты; N_j – число вейвлет-коэффициентов для уровня j; $\eta_j = \log_2\left(\frac{1}{N_j}\sum_k |d_x(j,k)|^2\right)$.

Для построения доверительного интервала $\hat{H}(j_1, j_2)$ необходимо оценить дисперсию оценки H, которая в работе [43] выражается формулой:

$$\sigma_H^2 = \mathrm{var}[\hat{H}(j_1, j_2)] = \frac{2}{N_{j1}\ln^2(2)} \frac{1-2^J}{1-2^{-J-1}(J^2+4)+2^{-2J}}, \quad (47)$$

где $J = j_2 - j_1$, $N_{j1} = 2^{-j1}n$ – число коэффициентов уровня j_1.

В результате можем записать доверительный интервал $\hat{H}$:

$$\hat{H} - \sigma_H z_\beta \le H \le \hat{H} + \sigma_H z_\beta, \quad (48)$$

где z_β – квантиль $1-\beta$ стандартного нормального распределения.

В качестве материнского вейвлета предлагается (Abry) [44] использовать вейвлеты “daubechies”. Выбор такого типа вейвлета подтверждается во многих исследованиях, например [46].

Кроме рассмотренных выше, существуют иные методы оценки, например, так называемый индекс фрактальности.

Наиболее совершенными методами оценки H являются методы, основанные на спектральных свойствах. Они являются наиболее точными в смысле несмещённости оценки на всём участке H от 0 до 1, а также позволяют определить свойства оценки.

Рассчитаем показатель Херста по ряду методов. Далее отберем активы в портфель с наибольшим значением полученного показателя. Портфель построим по прогнозным значениям доходностей отобранных активов, более подробно методика изложена в работе [1], проверим доходность портфелей на упреждающем периоде. Результаты расчетов представим в таблице. Также построим портфель со случайным отбором, сравним результаты.

Отрицательное значение доли акций в портфеле наблюдается, т.к. разрешены «короткие продажи». Суть «коротких продаж» можно выразить следующим образом: инвестор продает ценные бумаги по действующей в настоящий момент времени рыночной цене и остается с открытой позицией. Он выжидает удобный момент, чтобы купить их обратно, но более низкой цене, чтобы закрыть позицию. Если продавец впоследствии сможет выкупить указанные акции по более низкой цене, то он получит прибыль, при росте цен – понесет убыток.

Таблица

Структура и поступреждающая доходность портфелей

Случайный отбор активов		Метод нормированного размаха		Метод модулей приращений	
Акции	Доля в портфеле	Акции	Доля в портфеле	Акции	Доля в портфеле
Сургутнефтегаз	0,742	Норильский Никель	0,305	Татнефть	0,8960
Норильский Никель	0,591	Лукойл	-1,152	Северсталь	0,0230
Газпром	-0,001	Газпром	2,206	Русс-Инвест	0,4700
Лукойл	-1,047	Сбербанк	1,049	Русгидро	0,0008
МТС	0,565	Роснефть	-1,659	МТС	-0,0550
Сбербанк	0,151	Сургутнефтегаз	0,252	Лукойл	-0,3090
ПОД	-0,4855	ПОД	1,04	ПОД	0,21

Из таблицы видно, что портфели с предварительным отбором дают положительную доходность на упреждающем отрезке времени (15 дней). Использование метода нормированного размаха для отбора позволяет построить портфель с наибольшей доходностью. Таким образом, можно сказать, что отбор активов с помощью показателя Херста приводит к повышению эффективности инвестирования.

Список источников

1. Касаткин С.Е. Формирование портфеля ценных бумаг из акций с персистентной динамикой / С.Е. Касаткин, Т.В. Куликова // Современная экономика: проблемы и решения. – 2011. – №2(14). – С.143-151.

2. Предпосылки введения количественных мер эффективности для ГЭР / Архипов В.М., Захаров И. Ю., Науменко В.В., Смирнов С.Н.: Препринт WP16/2007/05. – М.: ГУ ВШЭ, 2007. – 40 с.

3. Kirichenko L, Radivilova T. Comparative analysis of statistical properties of the hurst exponent estimates obtained by different methods / Information Models of Knowledge / Krassimir Markov, Vitalii Velychko, Oleksy Voloshin / Kiev, Ukraine – Sofia, Bulgaria, 2010.

4. Pesaran, Hashem M. Market efficiency today. IEPR working paper 05.41 (Dec., 2005), pp. 6–9.

5. Clegg R.G. A practical guide to measuring the Hurst parameter. R. G. Clegg. Computing science technical report (№ CS–TR–916), 2005.

6. Beran. Statistics for long-memory processes. Chapman and Hall, 1994.

7. Taqqu M.S., Teverovsky V., Willinger W. Estimators for long-range dependence: an empirical study. Fractals, 3(4):785–788, 1995

8. Taqqu M.S. and Teverovsky V. Robustness of whittle type estimators for time

series with longrange dependence. Stochastic Models, 13:723–757, 1997

9. Bardet J.-M., Lang G., Oppenheim G., Phillipe A., Stoev S., and Taqqu M.S.. Semi-parametric estimation of the long-range dependence parameter: A survey. In P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors, Theory and applications of long-range dependence, pages 557–577. Birkhauser ¨ , 2003

10. Sheinkman J.A., LeBaron B., Nonlinear dynamics and stock returns. Journal of business 62, 1989.

11. Cootner P. Comments on the variation of certain speculative prices. Cambridge: MIT Press, 1964a.

12. Rose O. Estimation of the Hurst parameter long-range dependent time series. Institute of Computer Science, University of Wurzburg, Feb 1996.

13. Taqqu M.S. Self-similar processes. In Encyclopedia of Statistical Sciences 8, 352-357. Wiley, New York, 1988.

14. Kolmogorov A.N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. C.R. Acad. Sci. USSR (N.S.) 30 301-305, 1941.

15. Mandelbrot B.B. Self-similar error clusters in communication systems and the concept of conditional stationary. IEEE Transition on Communication Technology COM – 13 71-90, 1965.

16. Mandelbrot B.B. Some noise with 1/f spectrum, a bridge between direct current and white noise, IEEE Trans. Inform. Theory 13 289-298, 1967.

17. Mandelbrot B.B. Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. Internat. Econom. Rev. 10 82-113, 1969.

18. Mandelbrot B.B. A fast fractional Gaussian noise generator. Water Resourses Research, 7 543-553, 1971.

19. Mandelbrot B.B., Taqqu M.S. Robust R/S analysis of long run serial correlation. In Proceedings of the 42nd Session of the ISI 48 (Book 2) 69-99. Internat. Statist. Inst., Tokyo. 1979.

20. Mandelbrot B.B., Van Ness J.W. Fractional Brownian motion, fractional noises and application. SIAM Rev. 10 422-437, 1968.

21. Mandelbrot B.B., Wallis J.R. Some long-run properties of geophysical records. Water Recourses Research 5 321-340, 1969.

22. Beran J. Statistical methods for data with long-range dependence. Statistical Science 1992, Vol. 7, No. 4, 404-427.

23. Willinger W., Taqqu M.S., Leland W.E., Wilson D.V. Self-similarity in high-speed packet traffic: analysis and modeling of ethernet traffic measurements. Statistical science Vol. 10, N1, 67-85, 1995.

24. Granger C.W., Joyeux R. An introduction to long-memory time series models and fractional differencing. J.Time ser. Anal. 1 15-29, 1980.

25. Hosking, J.R. Fractional differencing. Biometrika 68 165-176, 1981.

26. Whittle P. Estimation and information in stationary time series. Arkiv for Matematiik, 2(23):423-434, 1953.

27. Dahlhaus R. Efficient parameter estimation for self-similar processes. The annals of Statistics, 17(4):1749-1766, 1989.

28. Fox R., Taqqu M.S. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. The Annals of Statistics, 14(2):517-532, 1986.

29. Geweke J., Porter-Hudak, S. The estimation and application of long memory time series models. J. Time Ser. Anal. 4 221-237, 1983.

30. Robinson P.M. Semiparametric analysis of long-memory time series. Ann. Statist. 22 515-539.

31. Beran J. Statistics for long-memory processes. Chapman & Hall, 1994.

32. Rama C. Long range dependence in financial markets, 2005.

33. Hurst H.E. Long-term storage capacity of reservoirs. Transactions of the American Society of Cicil Engineers 116 770-779, 1951.

34. Mandelbrot B.B. Limit theorems on the self-normalized range of weakly and strongly dependent processes. Zeitschrift fur Wahrscheinlichkeitstheoie und verwandte Gebiete, vol. 31, pp. 271-285, 1975.

35. Петерс Э. Фрактальный анализ финансовых рынков: применение теории хаоса в инвестициях и экономике. – М.: Интернет-трейдинг, 2004.

36. Sun W., Rachev S., Fabozzi F.J. Long-range dependence, fractal processes, and intra-day data. International Handbooks on Information Systems, II, 543-585, 2008.

37. Taqqu, M. S., and V. Teverovsky Estimating long-range dependence in finite and infinite variance sereis, in R. Adler, R. Feldman, and M. S. Taqqu, eds., A Practical Guide to Heavy Tails, Birkh¨auser: Boston, 1998.

38. Lo, A. W. Long-term memory in stock market prices. Econometrica, vol. 59, pp. 1279-1313, 1991.

39. Teverovsky V., Willinger W., A critical look at Lo's modified R/S statistic., Journal of Empirical Finance, vol. 9, pp. 455-474, 1999.

40. Abry P. Self-similarity and long-range dependence through the wavelet lens. P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch. Theory and applications of long-range dependence, Birkhäuser, 2003.

41. Graf H.P. Long-range correlations and estimation of the self- similarity parameter. PhD thesis, ETH Zurich. 1983.

42. Taqqu M.S. Fractional Brownian motion and long-range dependence. In P. Doukhan, G. Oppenheim, M.S. Taqqu, editors, Long-range Dependence: Theory and applications. Birkhauser, 2001.

43. Wornell G. W., "Wavelet-based representations for the 1/f family of fractal processes", Proceedings of the IEEE, vol. 81, n. 10, October 1993, pp.1428-1450

44. Abry P. Wavelet analysis of long-range dependent traffic. P. Abry, D. Veitch. IEEE/ACM Transactions Information Theory № 1(44), 1998.

45. Daubechies I., Ten Lectures on Wavelets. Philadelphia, PA: SIAM, 1992.

46. Ciftlikli C., Gezer A. Comparison of Daubechies wavelets for Hurst parameter estimation, 2010.

47. Cox D., Lewis P., "The Statistical Analysis of Series of Events," London: Chapman & Hall, 1966.

48. Cheridito P. Gaussian moving averages, semimartingales and option pricing, Stochastic Process. Appl., 109, pp. 47–68, 2004.

49. Mandelbrot B.B. Fractals and Scaling in Finance: Discontinuity, Concentration, Risk., Springer, New York, 1997.

50. Rea W, Oxley L, Reale M, Brown J . Estimators for long range dependence: an empirical study. Electronic Journal of Statistics. 2009.

HURST COEFFICIENT: CALCULATION METHOD AND POSSIBILITIES OF USE IN PROBLEMS OF PORTFOLIO INVESTMENT

Borovikov Ilya Mikhaylovich,
Post-graduate student of Institute of Management, Marketing and Finances (Voronezh); ilyaflash@list.ru;
Kulikova Tatyana Vyacheslavovna,
Post-graduate student of the Chair of Information Technolo-gies and Mathematical Methods in Economy of Voronezh State University; kulikova-tanya08@mail.ru;
Tinyakova Viktoriya Ivanovna,
Dr. Sc. of Economy, Professor of the Chair of Information Tech-nolo-gies and Mathematical Methods in Economy of Voronezh State University; tvikto-ria@yandex.ru

It is offered to form initial structure of the portfolio securities from shares that have the greatest value of Hurst coefficient. The methods of calculating of this indicator are considered.
Keywords: portfolio of assets, initial structure of the portfolio, Hurst coefficient.